

Cognitive Architectures for Language Agents

Year: 2023 | Citations: 265 | Authors: T. Summers, Shunyu Yao, Karthik Narasimhan, Thomas L. Griffiths

Abstract

Recent efforts have augmented large language models (LLMs) with external resources (e.g., the Internet) or internal control flows (e.g., prompt chaining) for tasks requiring grounding or reasoning, leading to a new class of language agents. While these agents have achieved substantial empirical success, we lack a systematic framework to organize existing agents and plan future developments. In this paper, we draw on the rich history of cognitive science and symbolic artificial intelligence to propose Cognitive Architectures for Language Agents (CoALA). CoALA describes a language agent with modular memory components, a structured action space to interact with internal memory and external environments, and a generalized decision-making process to choose actions. We use CoALA to retrospectively survey and organize a large body of recent work, and prospectively identify actionable directions towards more capable agents. Taken together, CoALA contextualizes today's language agents within the broader history of AI and outlines a path towards language-based general intelligence.